Geometric Object/Image Relations for RADAR

D. Gregory Arnold[a] and Kirk Sturtz[b]

[a]Air Force Research Lab, AFRL/SNAT, Bldg. 620, 2241 Avionics Circle, WPAFB, Ohio 45433-7321

[b]Veridian Incorporated, 5200 Springfield Pike, Suite 200, Dayton, Ohio 45431

ABSTRACT

Recent research in invariant theory has determined the fundamental geometric relation between objects and their corresponding "images". This relation is independent of the sensor (ex. RADAR) parameters and the transformations of the object. This relationship can be used to extract *3-D* models from image sequences. This capability is extremely useful for target recognition, image sequence compression, understanding, indexing, interpolating, and other applications.

Object/image relations have been discovered for different sensors by different researchers. This paper presents an intuitive form of the object/image relations for RADAR systems with the goal of enhancing interpretation.

This paper presents a high level example of how a *3-D* model is constructed directly from RADAR (or SAR) sequences (with or without independent motion). The primary focus is to provide a basic understanding of how this result can be exploited to advance research in many applications.

Keywords: RADAR ATR, geometric invariants, Lie group analysis, object/image relations, standard position method

1. INTRODUCTION

Recent research has yielded the fundamental geometric relation between objects and "images" (for RADAR, SAR, UHRR, EO, and IR sensors).[1–4] This paper provides a high level demonstration of the power of these results for analyzing RADAR data and attempts to develop some intuition into the nature of these relations.

We discuss several applications of the object/image relation including extracting a *3-D* model from an image sequence (with known correspondences) using only the assumption that the object undergoes a rigid transformation. This algorithm is based on point scatterers, however line features and surfaces could also be used with the additional relation found in.[5] We do not elaborate upon the correspondence problem, but assume optical flow-like techniques can be used to accurately and efficiently track points over short image sequences.[1]

The advantage of using this object/image relation, which employs *3-D* and *1-D* invariants, is that it circumvents estimating sensor and object orientation parameters. Once the ATR or *3-D* model reconstruction is complete, then these nuisance parameters can be estimated efficiently.

A *3-D* model will be referred to as an '*object*', and a *1-D* projection of that object (with a RADAR at an arbitrary orientation) as an '*image*.'

2. THE SENSOR MODELS

The sensor model characterizes the transformation group acting on the object, and the projection from $\Re^3$ to $\Re^1$. The model presented here is valid for a set of point scattering centers undergoing the same rigid motion. The theory will support more general models, but this is a good starting point.

We assume that the object is in the far-field of the RADAR. Therefore, geometrically, the RADAR is simply an orthographic projection onto a line (where the position on the line indicates relative range with respect to the RADAR). The line can be represented by a direction vector (the RADAR look direction) and is always normal (unity magnitude).

E-mail: garnold@mbvlab.wpafb.af.mil and ksturtz@mbvlab.wpafb.af.mil

2.1. The *2-D* to *1-D* Orthographic Projection

We will first develop a case analogous to RADAR, *2-D* orthographic projection onto *1-D* . This case is nice because it contains all the essential components of the RADAR case, but is fundamentally easier to visualize and relate due to the reduction in dimensionality. The relation between the measured range, r, and the feature location, $\{x, y\}$, can be expressed as

$$r = \begin{bmatrix} \rho_1 & \rho_2 & 0 \end{bmatrix} E \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

where

$$E = \begin{bmatrix} & R & a \\ & & b \\ 0 & 0 & 1 \end{bmatrix}$$

and

$$\rho_1^2 + \rho_2^2 = 1$$

such that $R \in \mathrm{SO}(2)$ (rotations), and $\{a, b\}$ is a vector denoting a rigid translation of $\Re^2$. Thus the Euclidean group defines the transformations acting on the modeled objects, and the projection is given by ρ_1 and ρ_2, yielding the range, r.

2.2. The *3-D* to *1-D* Orthographic Projection (RADAR)

The *3-D* case is exactly the same, except one extra dimension must be added to the model. The relation between the measured range of the feature location is

$$r = \begin{bmatrix} \rho_1 & \rho_2 & \rho_3 & 0 \end{bmatrix} E \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

where

$$E = \begin{bmatrix} & & & a \\ & R & & b \\ & & & c \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$\rho_1^2 + \rho_2^2 + \rho_3^2 = 1$$

such that $R \in \mathrm{SO}(3)$ (rotations), and $\{a, b, c\}$ is a vector denoting a rigid translation of $\Re^3$. Finally, $\{\rho_1, \rho_2, \rho_3\}$ represent the RADAR look direction.

3. INVARIANTS AND THE OBJECT/IMAGE RELATION

The fundamental object/image relation expresses a geometric relation (constraint) between a *3-D* object and its image. The particular invariants and number of points required for an invariant depend upon the transformation group associated with the sensor model.

Although constructing invariants is difficult, once an invariant has been found, it is typically simple to compute. Object/image (O/I) relations are an application of invariant theory as applied to the world viewed by a sensor (RADAR in this case). O/I relations provide a formal way of asking, “what are all possible images of this object,” and “what are all possible objects that could produce this image?” Clearly, this is a very powerful formalism and it is well suited to ATR.

Fundamentally the O/I relations can be viewed as the result of elimination of the unknown parameters in the model describing the projection of the 3-D world onto the sensor. Thus, the RADAR O/I relation can be derived by eliminating the group parameters associated with target motion, and the parameters associated with the RADAR look direction. The result is an equation relating the object to its RADAR scatterer range (projected image) written in terms of their associated invariants. More specifically, the equation relates *3-D* Euclidean invariants (inter-scatterer

3-D distances, determinants, or inner products) to translation invariant *1-D* inter-scatterer range differences in the *1-D* backscattered RADAR signal. The basic theory states that given 6 *different 1-D* RADAR images and a set of 4 (generally non-coplanar) scatterers, the *3-D* locations of the scatters can be recovered completely as a function of time.

3.1. The *2-D* to *1-D* Orthographic Projection Case

An invariant of *2-D* objects (undergoing Euclidean transformations) requires a minimum of three points, $P_i = \{X_i, Y_i, 1\}$. Let the point P_i correspond to the i^{th} column of a 3×3 matrix. By translating and rotating appropriately, one can always put the matrix into the standard position,

$$\begin{bmatrix} 0 & I_1 & I_2 \\ 0 & 0 & I_3 \\ 1 & 1 & 1 \end{bmatrix}$$

where $\{I_1, I_2, I_3\}$ are invariant under *2-D* Euclidean transformations. Standard linear algebra proves that a unique transformation exists to make this change of basis. It is not obvious here without further explanation, but the invariants are functions of the Euclidean distance, determinants, and inner products.

The three *1-D* points, $q_i = \{r_i\}$, can always be transformed to the standard position,

$$\begin{bmatrix} 0 & i_1 & i_2 \end{bmatrix}.$$

The resulting expressions for i_1 and i_2 are invariant with respect to translation. Specifically, $i_1 = r_2 - r_1$, and $i_2 = r_3 - r_1$.

The fundamental object/image relation can now be determined by solving the model projection equations with respect to the unknowns, ρ_1 and ρ_2,

$$\begin{aligned} i_1 &= I_1 \rho_1 \\ i_2 &= I_2 \rho_1 + I_3 \rho_2 \end{aligned}$$

and substituting into the unity vector constraint,

$$\rho_1^2 + \rho_2^2 = 1.$$

After some simplification, the resulting object / image relation is

$$\begin{aligned} \left(\frac{i_1}{I_1}\right)^2 + \left(\frac{i_1 I_2 - I_1 i_2}{I_1 I_3}\right)^2 &= 1 \\ \left(\frac{I_2^2 + I_3^2}{I_1^2 I_3^2}\right) i_1^2 + \left(\frac{-2 I_2}{I_1 I_3^2}\right) i_1 i_2 + \left(\frac{1}{I_3^2}\right) i_2^2 &= 1 \\ \left(I_2^2 + I_3^2\right) i_1^2 + \left(-2 I_1 I_2\right) i_1 i_2 + \left(I_1^2\right) i_2^2 &= I_1^2 I_3^2 \end{aligned}$$

This particular formulation assumes that the 3 points are distinct and not co-linear, however the O/I relation still holds if this is not the case. If any two points are the same, then the equation becomes trivial, $0 = 0$, which is not very interesting. If all three points are co-linear, then the equation reduces to $I_2 i_1 = I_1 i_2$ or $\frac{i_1}{i_2} = \frac{I_1}{I_2}$ which are the standard *1-D* invariants one would use for co-linear points. In other words, this object/image relation is truly a generalization of the special cases that have been well studied over the last 10 years. Finally, the intermediate form is useful since it clearly shows the elliptic nature of this equation with respect to the image invariants.

Stiller[3] has shown these geometric constraints are satisfied *if and only if* the image, which determines $\{i_j\}_{j=1}^2$, can be formed by the object, which determines $\{I_j\}_{j=1}^3$.

3.2. The *3-D* to *1-D* Orthographic Projection Case

RADAR is essentially an orthographic projection from *3-D* to *1-D* in the far-field. An invariant of *3-D* objects (undergoing Euclidean transformations) requires a minimum of four points, $P_i = \{X_i, Y_i, Z_i, 1\}$. Let the point P_i correspond to the i^{th} column of a 4×4 matrix. By translating and rotating appropriately, one can always put the matrix into the standard position,

$$\begin{bmatrix} 0 & I_1 & I_2 & I_3 \\ 0 & 0 & I_4 & I_5 \\ 0 & 0 & 0 & I_6 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

where $\{I_1, I_2, I_3, I_4, I_5, I_6\}$ are invariant under *3-D* Euclidean transformations. Standard linear algebra proves that a unique transformation exists to make this change of basis. As noted previously for the *2-D* case, the invariants are functions of the Euclidean distance, determinants, and inner products.

The four *1-D* projections, $q_i = \{r_i\}$, can always be transformed to the standard position,

$$\begin{bmatrix} 0 & i_1 & i_2 & i_3 \end{bmatrix}.$$

The resulting expressions for i_1, i_2, and i_3 are invariant with respect to translation. Specifically, $i_1 = r_2 - r_1$, $i_2 = r_3 - r_1$, and $i_3 = r_4 - r_1$.

The fundamental object/image relation can now be determined by solving the model projection equations with respect to the unknowns, ρ_1, ρ_2, and ρ_3,

$$\begin{aligned} i_1 &= I_1\rho_1 \\ i_2 &= I_2\rho_1 + I_4\rho_2 \\ i_3 &= I_3\rho_1 + I_5\rho_2 + I_6\rho_3 \end{aligned}$$

yielding

$$\begin{aligned} \rho_1 &= \frac{i_1}{I_1} \\ \rho_2 &= \frac{I_1 i_2 - I_2 i_1}{I_1 I_4} \\ \rho_3 &= \frac{I_1 I_4 i_3 - I_3 I_4 i_1 - I_1 I_5 i_2 + I_2 I_5 i_1}{I_1 I_4 I_6} \end{aligned}$$

which will be substituted into the unity vector constraint,

$$\rho_1^2 + \rho_2^2 + \rho_3^2 = 1.$$

After some simplification, the resulting object / image relation is

$$((I_3 I_4 - I_2 I_5)i_1 + I_1 I_5 i_2 - I_1 I_4 i_3)^2 + \left((I_2 i_1 - I_1 i_2)^2 + \left(i_1^2 - I_1^2\right) I_4^2\right) I_6^2 = 0.$$

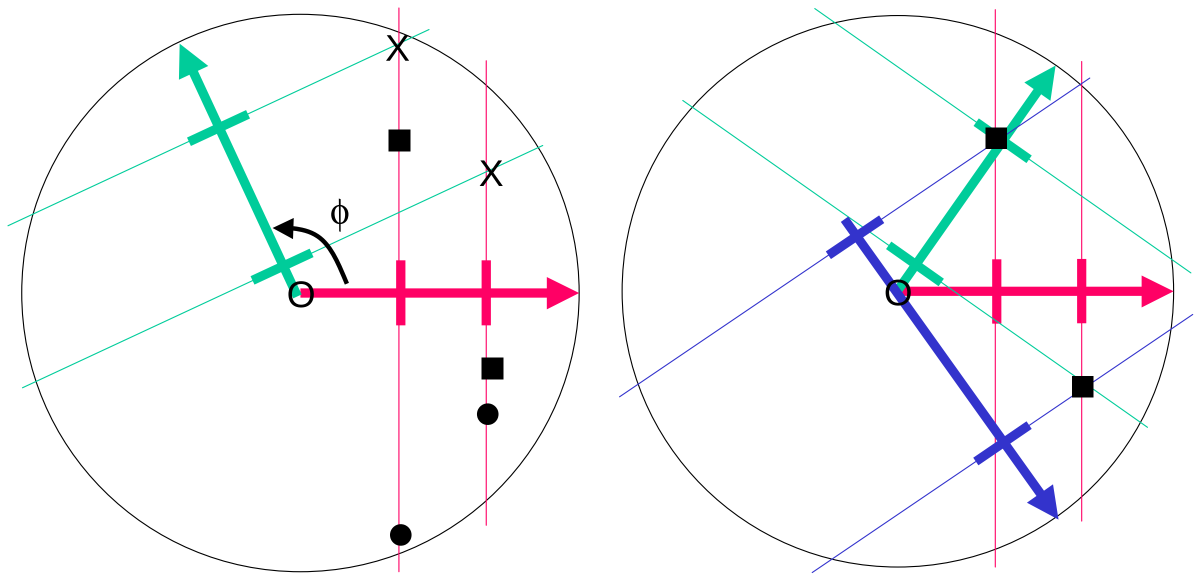
This particular formulation assumes that the 4 points are distinct and not co-planar. But, as before, the O/I relation still holds in the special cases of non-distinct, co-linear, and co-planar configurations. This O/I relation can be re-written to show that it is fundamentally an ellipsoid centered at the origin with respect to the image invariants.

Stiller[3] has shown these geometric constraints are satisfied *if and only if* the image, which determines $\{i_j\}_{j=1}^3$, can be formed by the object, which determines $\{I_j\}_{j=1}^6$.

4. VISUALIZATION OF THE OBJECT / IMAGE RELATIONS

Object / image relations provide a formal way of asking, “what are all possible images of this object,” and “what are all possible objects that could produce this image?” Clearly, this is a very powerful formalism and it is well suited to ATR.

The geometric relation between all objects and images is clearly spelled out in the object / image relation. This section attempts to provide a geometric insight into this relation to generate the appropriate intuition for understanding these results. We will examine the O/I relation for the *2-D* to *1-D* orthographic projection. This case directly generalizes to the true RADAR case, but is simpler to manipulate and visualize.



(a) (b)

Figure 1. An example of 'all possible objects given an image' (2-D orthographic projection). Three scattering centers are projected onto a line. The first scattering center is placed at the origin. The radius of the circle is defined by the range gate of the RADAR because two points cannot be in the same image if they are farther apart than the range gate. (a) demonstrates two views of an object. The relative angle between the views is not known, so the X's, squares, and circles show the consistent object for three different relative angles. (b) demonstrates the result after three views are collected. The object/image relation guarantees at most one object is consistent with three views (up to reflection).

4.1. All Possible Objects of an Image

Restating the problem, "given an image (therefore we know i_1 and i_2), what is the set of objects (which corresponds to selecting I_1, I_2, and I_3) that satisfy the O/I relation?" An obvious way to display this is a *3-D* figure with I_1, I_2, and I_3 as coordinate axes (thus each point on the graph represents a solution to the O/I relation and thus a potential object).

Figure 1 shows another representation of all possible objects. This representation is dependent on the standard position chosen. The first scattering center is at the origin of the circle, and the radius of the circle corresponds to the length of the range gate used by the RADAR. The range gate provides an additional constraint on the I's by limiting the maximum Euclidean distance between the *3-D* points. Since we do not assume any information about the orientation of the object, the figure can only show the relative change in angle between different look directions. As this relative orientation changes between the look directions, the intersection defines objects consistent with both views. Finally, if a third view is added, at most one relative orientation exists between the three views that is consistent with all three images (figure 1(b)).

The first representation (using the invariants as the coordinate axes) makes visualization impossible when the number of invariants is greater than three. The second representation will always be in *2-D* or *3-D* depending on the model considered.

The obvious applications of knowing all possible objects of an image is object recognition or ATR. The intersection of a database of potential objects with all possible objects as defined by the image (or images) yields a very small list of potential matches. Huber[7] and Arnold[8] provide some initial insight into the complexity of this search problem.

Another useful application is constructing *3-D* models directly from a set of images. Stuff[1] and Arnold[9] demonstrate the necessary tools to do this. Each image has a set of different objects that could have produced it (a manifold defined by the solutions to the object/image relation). By intersecting the manifolds, the set of possible objects is thereby reduced until, with enough images, at most one object is consistent with all the images.

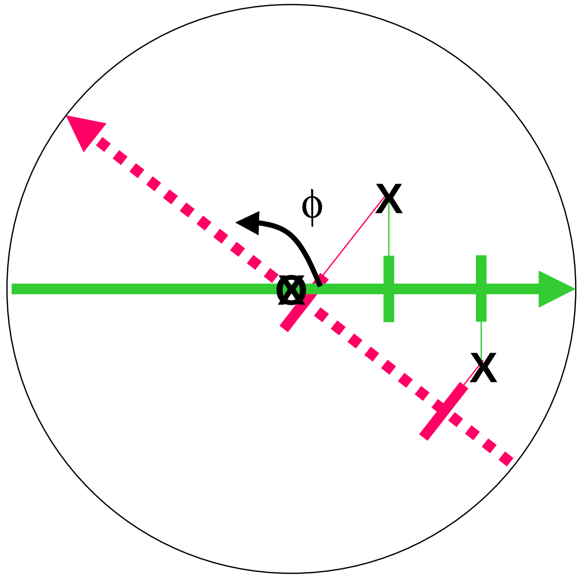


Figure 2. An example of 'all possible images given an object' (2-D orthographic projection). Three scattering centers are represented by the X's. The first scattering center is at the origin and similar to figure 1, the radius of the circle is defined by the range gate of the RADAR because two points cannot be in the same image if they are farther apart than the range gate. The relative angle between the sensor and the object is not known, so the two lines represent two possible images of the object.

4.2. All Possible Images of an Object

"What are all possible images of this object?" and "when (from what viewpoint) do these two objects look the same?" are fundamentally useful questions for evaluating the discrimination capability of an ATR system.

An obvious way to answer this question is a *2-D* figure with i_1 and i_2 as coordinate axes (thus each point on the graph represents a solution to the O/I relation and thus a consistent image). This figure is an ellipse for orthographic projection and the standard position presented here. However, as discussed above, this representation is not practical if the number of invariants is greater than three. Still, for the RADAR case, the appropriate representation is an ellipse. Therefore, another way to visualize the *3-D* model reconstruction problem is as an ellipse estimation problem. Algorithms that rapidly estimate the parameters of an ellipse in the presence of noise are essentially what is needed for reconstructing the models from orthographic projection.

Figure 2 shows the representation corresponding to that used for demonstrating 'all possible objects' in the previous section. Once again, the first scattering center is at the origin of the circle, and the radius of the circle corresponds to the length of the range gate used by the RADAR.

5. GENERAL PROCESSING ALGORITHM

Figure 3 shows a general object recognition and tracking algorithm based on the fundamental object/image relation presented in this paper. Each box can be discussed in terms of invariants and the fundamental object/image relation.

1. *Image Acquisition*

   Image acquisition refers to obtaining the raw data from the imaging sensor. This paper discusses the general RADAR case (orthographic projection). Analogous results to this paper have been presented for weak perspective and full perspective sensors.[2–4]

2. *Feature Detection / Labeling*

   Extract features from the imagery and label different features types (corner, plate, curved, etc. analogous to points, lines, surfaces, colors of visible imagery). Ideally, the feature detectors are based on photometric (intensity) invariants.

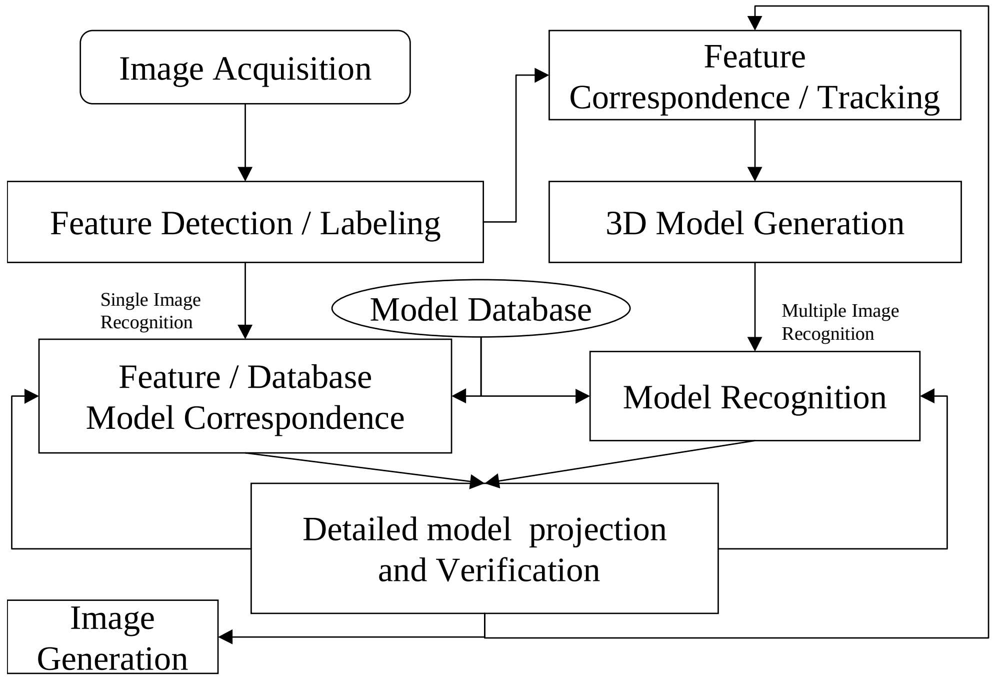


Figure 3. General object recognition and tracking algorithm based on the fundamental object/image relation discussed in this paper.

3. *Feature / Database Model Correspondence*

   This is the single view recognition problem. This is an application of the object/image relation using a model database similar to those presented by Weiss[4] and Huber.[7]

4. *Feature Correspondence / Tracking*

   Scattering centers can be tracked over short time periods[1] and used to construct a *3-D* model. This *3-D* model can then be used to improve / predict future feature locations. Initially, many possible objects may exist, but as the sequence progresses, the choices will consolidate or disappear.

5. *3-D Model Generation*

   As discussed in section 4.1 and 4.2, construction of the *3-D* model has a straightforward interpretation of intersecting manifolds, fitting an ellipse, or solving a system of over-determined equations.

6. *Model Recognition*

   Once a *3-D* (scattering center) model has been created, it can be compared against a database of *3-D* objects.

7. *Detailed model projection and verification*

   Once a model has been identified (based on certain features), the corresponding database model can be used to fill in missing, weak, local, and global features. It is essential to verify the object hypothesis by projecting the model into the image and performing a pixel level comparison!

8. *Image Generation*

   The sensor model can (generally) be estimated given an image and a known model. This sensor model estimate can then be used to re-project the model into novel views. This is very similar to detailed model projection and verification.

6. CONCLUSIONS

Object/image relations are a new tool to address many of the current problems in ATR. By applying this geometric constraint to the features, RADAR algorithms can be decomposed into feature correspondence, *3-D* model generation, and then any desired processing. The *3-D* model generation and desired processing are relatively simple processes in this general algorithm. Therefore, the capabilities of future systems are fundamentally limited by the capabilities of the feature correspondence algorithms. Image sequences are a particularly well suited application of this work because feature tracking over very short temporal changes is a much easier task than feature correspondence after arbitrary time lapses. Thus, this approach can be thought of as a generalization of the SAR image formation process. Future work includes finding additional types of invariants, developing invariants for non-rigid motion, and most importantly, examining the combinatorics of the feature correspondence problem.

REFERENCES

1. M. A. Stuff, "Three-dimensional analysis of moving target radar signals: Methods and implications for atr and feature aided tracking," in *Proceedings SPIE Int'l Conf.*, (Orlando, FL), Apr 1999.
2. P. Stiller, C. Asmuth, and C. Wan, "Invariants, indexing, and single view recognition," in *Proc. ARPA Image Understanding Workshop*, pp. 1432–1428, (Monterrey, CA), Nov 13-16 1994.
3. P. Stiller, C. Asmuth, and C. Wan, "Single view recognition - the perspective case," in *Proceedings SPIE Int'l Conf., Vision Geometry V*, vol. 2826, pp. 226–235, (Denver, CO), Aug 1996.
4. I. Weiss, "Model-based recognition of 3d objects from one view," in *Proc. DARPA Image Understanding Workshop*, pp. 641–652, Morgan Kauffman, (Monterey, CA), Nov 20-23 1998.
5. P. Stiller, "Object recognition via configurations of lines," in *Proceedings SPIE Int'l Conf., Vision Geometry VII*, vol. 3454, pp. 76–86, (San Diego, CA), Aug 1998.
6. G. Arnold, K. Sturtz, and V. Velten, "Lie group analysis in object recognition," in *Proc. DARPA Image Understanding Workshop*, Morgan Kauffman, (New Orleans, LA), May 1997.
7. B. Huber, P. Stiller, C. Wan, and T. Shah, "Invariant geometric hashing," in *Proc. of the 27th Workshop on Applied Imagery and Pattern Recognition*, Oct 1998.
8. D. G. Arnold and K. Sturtz, "Complexity analysis of atr algorithms based on invariants," in *Proceedings Computer Vision Beyond the Visible Spectrum (CVBVS)*, IEEE Computer Society, (Hilton Head, SC), June 2000. To Appear.
9. G. Arnold, K. Sturtz, and G. Power, "Analyzing image sequences using invariants," in *CISST*, (Las Vegas, Nevada), June 1999.